

5-4-2018

Analyzing the Importance of Phosphorylation and Ubiquitination of Amino Acids in the Enhancer of Rudimentary Protein, ERH

Amber L. Voyer

Follow this and additional works at: <https://digitalcommons.brockport.edu/honors>

 Part of the [Biology Commons](#)

Analyzing the Importance of Phosphorylation and Ubiquitination of Amino Acids in the
Enhancer of Rudimentary Protein, ERH

A Senior Honors Thesis

Submitted in Partial Fulfillment of the Requirements
for Graduation in the Honors College

By
Amber Voyer
Biology & Medical Technology Major

The College at Brockport
May 4, 2018

Thesis Director: Dr. Stuart Tsubota, Professor, Biology

*Educational use of this paper is permitted for the purpose of providing future
students a model example of an Honors senior thesis project*

Table of Contents

ABSTRACT	4
INTRODUCTION.....	4
METHODOLOGY.....	6
1. CREATION OF MUTANT HUMAN CODING REGIONS OF <i>E(R)</i>	6
2. CONSTRUCTION OF RECOMBINANT VECTORS	9
3. INSERTION OF THE MUTATED HUMAN <i>E(R)</i> GENE INTO THE FINAL VECTOR, PATTB	12
4. PHENOTYPIC ANALYSIS OF MUTANT <i>DROSOPHILA MELANOGASTER</i> USING GENETIC MANIPULATION	13
RESULTS.....	15
1. EVIDENCE OF THE SUCCESSFUL CREATION OF RECOMBINANT VECTORS	15
2. RESULTS OF THE CROSSES OF MUTANT FEMALES TO MALES CONTAINING THE DESIGNED TRANSGENES	16
DISCUSSION	19
CONCLUSION.....	20
APPENDICES	21
BIBLIOGRAPHY	23

Figures and Tables

FIGURE 1. MUTATIONS STUDIED	7
FIGURE 2. LOCATION OF PHOSPHORYLATION AND UBIQUITINATION SITES IN CRYSTAL STRUCTURE	7
FIGURE 3. AMINO ACID SUBSTITUTIONS FOR PHOSPHORYLATION	8
FIGURE 4. AMINO ACID SUBSTITUTION FOR UBIQUITINATION	9
FIGURE 5. NCOI AND PciI RESTRICTION ENZYME SITES	10
FIGURE 6. DIGEST FOR DETERMINING INSERT ORIENTATION	11
FIGURE 7. pSMART-EMPTY E(R) AND CODING REGION INSERT	12
FIGURE 8. PATTB VECTOR WITH E(R) GENE INSERT	13
FIGURE 9. CROSS OF HETEROZYGOUS FEMALE WITH E(R) DELETION AND MALE WITH E(R) TRANSGENE	14
FIGURE 10. GEL ELECTROPHORESIS OF pSMART VECTOR WITH CODING REGION INSERT	15
FIGURE 11. GEL ELECTROPHORESIS OF ISOLATED E(R) GENE	16
FIGURE 12. DIGEST OF PATTB PLASMID AND E(R) GENE INSERT FOR VERIFICATION	16
TABLE 1. THE E(R) ²⁷⁻¹ MUTATION RESULTS	17
TABLE 2. THE E(R) ²⁷⁻¹ R ^{HD1} DOUBLE MUTANT RESULTS	18
TABLE 3. THE E(R) ³⁷⁻⁶ N ^{ND-P} DOUBLE MUTANT RESULTS	19

Abstract

The *enhancer of rudimentary* gene, *e(r)*, has been found to possess certain amino acids with the potential for phosphorylation or ubiquitination. All of these amino acids are highly conserved among species. It is possible that any of these sites might play a role in Enhancer of Rudimentary Homolog protein, ERH, activation by their respective modifications. This possibility has been analyzed through the creation of transgenes with codon changes at the identified sites intended to either prevent or mimic phosphorylation or ubiquitination. Upon insertion of the *e(r)* transgenes into the *Drosophila melanogaster* genome, they have been crossed with two lethal and one low viability stock of flies. The location of the *e(r)* gene on the X chromosome allows basic phenotypic observations based on eye color to reveal whether any mutations exhibit wild-type ratios of males to females. It has been determined that the phosphorylation of Y19, Y22, S47, and Y92 are not necessary in the activation of ERH. Likewise, the prevention of ubiquitination at K41 and K90 does not effect protein function. A double amino acid change at T18 and S24 in the same transgene has resulted in low viability when compared to the wild-type, suggesting a decreased amount of ERH activity as a result of this mutation in regards to the *Notch* signaling pathway.

Introduction

The *enhancer of rudimentary* gene, *e(r)*, encodes the protein Enhancer of Rudimentary Homolog, ERH. This protein is found in many different species such as *Homo sapiens* (human), *Drosophila melanogaster* (fruit fly), *Mus musculus* (mouse), and *Arabidopsis thaliana* (flowering weed) (Gelsthorpe *et al.* 1997). Research is still being done to discover all of the activities that ERH may be involved in. However, the most notable action of ERH that has been identified so far has been the promotion of the progression and growth of cancer cells in humans. Breast cancer cells lines in particular have been shown to possess increased levels of ERH (Zafrakas *et al.* 2008). A similar over expression of the protein has also been shown in the testis and ovaries of humans. One of the ways that ERH may specifically aid in cell division within tumors is by regulating pyrimidine synthesis (Wojcik *et al.* 1994). This is an essential process for the replication of DNA and therefore cell division and growth.

It is evolutionarily significant to note that across species, vertebrates in particular, ERH is highly conserved (Gelsthorpe *et al.* 1997). This conservation lies in the function of the protein as well as the amino acid sequence itself. For example, in *D. melanogaster*, there has been

evidence of *e(r)* gene expression in cells undergoing DNA synthesis and replication. This is comparable to the role of *e(r)* gene expression in human cancer cells. There has also been evidence that ERH plays a role in the *Notch* signaling pathway of *D. melanogaster*. This pathway controls the development of undifferentiated cells and it requires the *e(r)* gene for proper expression. As a result of past studies, it can be said that *e(r)* positively regulates the *Notch* signaling pathway in *Drosophila* (Tsubota *et al.* 2011).

In terms of the conservation within the amino acid sequence of ERH, there is strong evidence that certain regions and sites are essential for proper function. When comparing the human and *Drosophila* versions of ERH, 76% of the sequence is identical (Fig. 1). By including different amino acids that are in the same position and have the same properties, the two proteins can be described as 84% similar. Relatively, this is a very high degree of conservation among two distinct species.

The human and *Drosophila* ERH are both 104 amino acids in length. The human protein has been analyzed by NetPhos. 3.1 (<http://www.cbs.dtu.dk/services/NetPhos/>) and NetphoSitePlus (www.phosphosite.org) and locations of possible phosphorylation or ubiquitination have been identified. The possible phosphorylation sites include T18, Y19, Y22, S24, S47, and Y92 while the ubiquitination sites may be K41 or K90 (Fig. 2). It is important to note that each of these sites are contained within the conserved portions of the protein. One exception is that there is a threonine in the *Drosophila e(r)* gene at the 47th position instead of a serine like in the human gene. Although this is not in a conserved portion of the gene, the two amino acids do have similar properties. It is common for serine and threonine to be interchangeable when it comes to the action of certain kinases. The conservation of the potential modification sites suggests that any of these amino acids could be important to proper protein function.

It is possible to study the human *e(r)* gene within *Drosophila* based on past research that has been done showing that the human ERH and *Drosophila* ERH are functionally equivalent (Tsubota *et al.* 2016). This means that a sick fruit fly with an *e(r)* deletion can be rescued with the insertion of the human *e(r)* gene. It is advantageous to study the actual human protein rather than a closely matched version of ERH because the results can be directly applied to *e(r)* function in the human body.

To study the hypothesis that one or more of these phosphorylation or ubiquitination sites requires modification for ERH to function properly, mutations within the *e(r)* gene were made. Some mutations were designed to prevent modification while others were made to mimic a modification, allowing the effects to be analyzed. After insertion of the mutant *e(r)* gene into *Drosophila*, crosses were done so that the resulting phenotypes could be assessed. The results will identify any sites that are required for normal ERH function and activation.

There are three mutant phenotypes that were used in order to determine which, if any, mutations can successfully rescue. The first is an *e(r)*²⁷⁻¹ single mutation that causes low viability. The mutation lacks the start of transcription as well as 43% of the coding region. The second is a lethal *e(r)*²⁷⁻¹ *r^{hd1}* double mutation. The *r^{hd1}* mutation causes the development of mutant wings. When combined with the *e(r)*²⁷⁻¹ deletion, the double mutant is lethal. The *e(r)*³⁷⁻⁶ *N^{nd-p}* double mutation is the third one tested. The *e(r)*³⁷⁻⁶ is similar to the *e(r)*²⁷⁻¹ as it also lacks the start of transcription. The combination of *e(r)*³⁷⁻⁶ with a weak *Notch* allele is lethal. The qualifications for “rescue” varies by each mutation. However, as all are X-linked alleles, a ratio of males to females possessing a copy of a designed *e(r)* mutation will reveal the effect of each.

Methodology

1. Creation of mutant human coding regions of *e(r)*

In order to begin the creation of recombinant vectors with mutations to be studied, mutant human *e(r)* coding regions were designed on paper (Appendix I). The DNA sequences are said to be “drosophilized” meaning that some of the amino acid codons were altered so that the most frequent codon for each amino acid in the *Drosophila* genome was used. Certain codons that are redundant for amino acids have been proven to be selected for more frequently in the DNA of certain species. For example, in humans it is more common to find aspartic acid coded for with the bases GAC. In contrast, *Drosophila* will most often use GAT to code for aspartic acid. By using the preferred nucleotides to code for amino acids, the ability of the *Drosophila* to synthesize the mutant protein was ensured but the polypeptide sequence was still consistent with that of the human ERH.

The nine mutants made were: S47A, S47E, Y92F, Y22F, Y19F, T18AS24A, T18ES24A, K90R, and K41R (Figs. 1 and 2). The first 7 mutants are testing phosphorylation sites and the last two are testing ubiquitination sites.

1 10 20 30 40 50 60 70 80 90 100

D. melanogaster wild-type:
MSHTILLVQPGARPETRT CDYESVNECMEGVCKIYEEHLKRRNPNTPTITYDISQLFDFIDTMVDISCMVYQKSTNTYAPYNKDWIKE I VLLRQAAFSSNT

Human wild-type:
MSHTILLVQPTKRPEGRRTADYESVNECMEGVCKMYEEHLKRMNPNPSITYDISQLFDFIDDLADLSCLVYRADTQTYQPNKDWIKE I VLLRQAQAGK

Phosphorylation Sites:
T18AS24A: A A
T18ES24A: E A
Y22F: F
S47A: A
S47E: E
Y92F: F
Y19F: F
Ubiquitination Sites:
K41R: R
K90R: R

Figure 1. Mutations Studied

The highlighted regions are conserved between the human and *D. melanogaster* wild-types. The location of each mutation can be seen in the wild-type sequences listed. The mutations and their positions are shown below the corresponding wild-type amino acids.

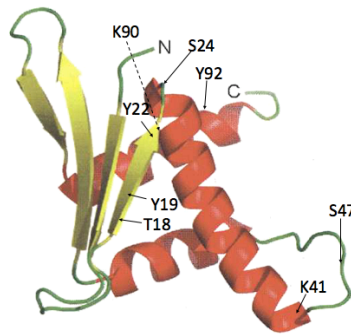


Figure 2. Location of Phosphorylation and Ubiquitination Sites in Crystal Structure

The crystal structure of the Human ERH with the location of the phosphorylation and ubiquitination sites labeled. This figure is a modification of a published figure (Arai et al. 2005).

One type of amino acid change in these mutants is designed to prevent modification. For example, positions with tyrosine in the wild-type are potential phosphorylation sites due to the presence of a hydroxyl group (see highlighted hydroxyl in Fig. 3A). To prevent the ability to phosphorylate at this site, a mutation is made by replacing tyrosine with phenylalanine. This choice has been made because phenylalanine and tyrosine only differ by the hydroxyl group. Otherwise, they are similar in size and aromatic properties (Fig. 3A). Following the same logic, serine and threonine can also be replaced with alanine which lacks the ability to be phosphorylated but maintains a similar size (Fig. 3B). The hydroxyl group capable of being phosphorylated in each is also highlighted in Fig. 3B.

The other type of amino acid change in this study is an amino acid phosphorylation mimic. This is able to test the possibility that modification may actually disrupt proper ERH

function. An example of this type of mutation would be a change from serine or threonine to glutamic acid. Serine and threonine, like tyrosine, have a hydroxyl group that allows phosphorylation. Glutamic acid is negatively charged, mimicking a negatively charged phosphoryl group on either a serine or threonine. Glutamic acid is also of comparable size to the wild-type amino acids when they are phosphorylated (Fig. 3C).

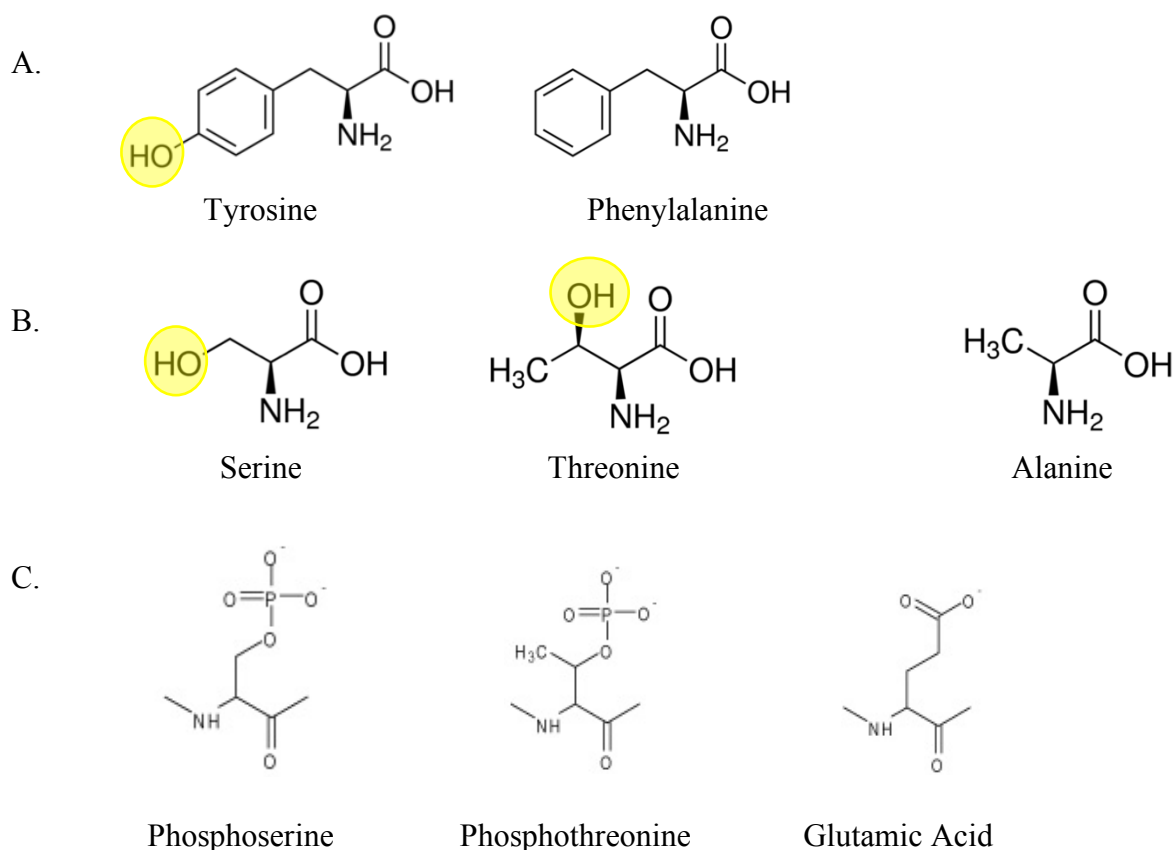


Figure 3. Amino Acid Substitutions for Phosphorylation

A) Phenylalanine is used to replace tyrosine to prevent the possibility of phosphorylation at the highlighted hydroxyl group. B) Alanine replaces serine or threonine in some of the mutants to prevent phosphorylation at one of the highlighted hydroxyl groups. C) Glutamic acid is used to mimic a phosphoserine or phosphothreonine.

The amino acid changes for the proposed ubiquitination sites are designed to prevent modification. The change is arginine substituted for lysine based on evidence that arginine is ubiquitination-resistant but resembles lysine in shape and charge (Fig. 4) (Xu and Jaffrey 2013).

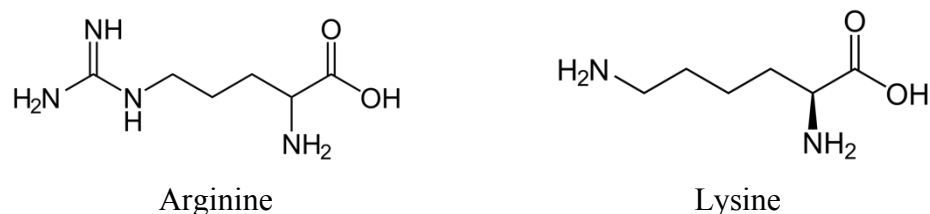


Figure 4. Amino Acid Substitution for Ubiquitination

The amino acid arginine which replaces lysine in order to prevent ubiquitination. They are similar in charge and size.

Two of the mutants created contain changes at two sites, 18 and 24. In the *Drosophila* ERH, these two positions have already been shown to be phosphorylated by casein kinase II. When both are prevented from phosphorylation, the protein is inactive. However, if only one is prevented, phosphorylation is possible and the protein is active. If both sites are replaced by a glutamic acid mimic, then ERH becomes inactive (Gelsthorpe *et al.* 2006). These results will be analyzed in the human ERH in the context of this study.

In designing the sequences to be sent to Integrated DNA Technologies, Inc. (<https://www.idtdna.com/site>) for synthesis, it was important to consider the species that the final vector would be inserted into. Every species has certain codons that they prefer for each amino acid and these triplets are used more frequently by the species. By using the preferred codons, it was more likely that the *Drosophila* would be successful in translating the mutated human *e(r)* gene after insertion (Vicario *et al.* 2007).

Another consideration was the ability to eventually create desired sticky ends on both sides of the synthesized double-stranded sequence of DNA for cloning the mutant sequences. Sticky ends are terminal single-stranded sequences that can base-pair with each other. It was decided to have a string of bases followed by a *PciI* restriction enzyme site, then the start codon ATG in the mutant DNA sequences. The other end of the sequence has the completion of the coding region followed by a *NcoI* site and then a few more bases. For both the *PciI* and *NcoI* sites that were included, the integrity of the sequence encoding the human ERH protein is maintained.

2. Construction of Recombinant Vectors

The synthesized mutant human *e(r)* coding regions flanked by *PciI* and *NcoI* sites were designed in such a way to allow their insertion into a vector called pSmart-empty *e(r)*. This

vector was created by adding the 5' untranslated region and 3' untranslated region without the coding region to a useful research vector called pSmart (Tsubota et al. 2016). At the end of the 5' untranslated portion, a required translation signal, CACC, is present. In between the two untranslated regions, instead of the coding region, is an NcoI site. This enables the insertion of any coding region desired, such as the human ERH coding region. Once the pSmart-empty *e(r)* vector was cut with NcoI, while maintaining the translation signal, the synthesized coding region was inserted into the proper place.

Recall that the synthesized mutant coding region has a NcoI site at one end and a PciI site at the other. Even though the vector was cut with NcoI, it is possible to insert both ends of the coding region into it, thereby making a circular vector. This is due to the fact that the sticky ends created by NcoI and PciI are able to base pair together (Fig. 5A and 5B). However, when this happens, the resulting sequence can no longer be recognized by either restriction enzyme. This is a helpful occurrence because the orientation of the coding region within the vector can be checked by locating the NcoI site and the NcoI/PciI site.

A.	NcoI Site	NcoI Sticky End	PciI Sticky End	PciI Site
	5'...CCATGG...3'	C CATGT	5'...ACATGT...3'
	3'...GGTACC...5'	GGTAC A	3'...TGTACA...5'
B.	NcoI/PciI			
	5'...CCATGT...3'			
	3'...GGTACA...5'			

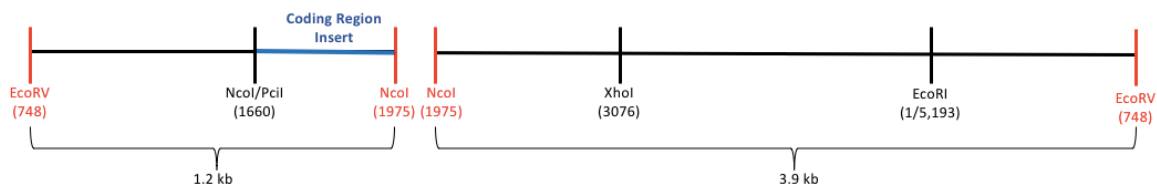
Figure 5. NcoI and PciI Restriction Enzyme Sites

A) The site that is recognized by the restriction enzyme NcoI and the sticky end that it creates. Likewise, the recognition site and resulting sticky end of PciI. B) The result of the combination of an NcoI sticky end and a PciI sticky end that is no longer recognized by either restriction enzyme.

It was important to verify the orientation of the coding region before proceeding to any purification steps. There is no way to predict which direction the coding region will insert as there is an equal chance for either way to occur. Therefore, a diagnostic digest was done in order to check the resulting fragment sizes. The positions of the NcoI/PciI site and NcoI/NcoI sites are found with an NcoI-EcoRV double digest. Depending on the fragment sizes, the recombinant

plasmid containing the desired orientation can be identified and used for the rest of the procedure (Fig. 6).

A. Desired Orientation



B. Incorrect Orientation

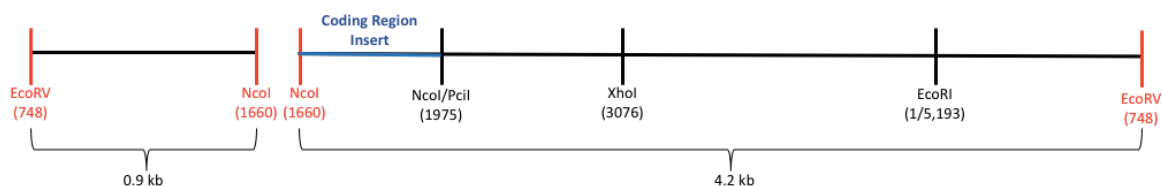


Figure 6. Digest for Determining Insert Orientation

A) The resulting fragments when the coding region inserts so that the 5' untranslated region is followed by the beginning of the coding region. B) The resulting fragment when the coding region inserts with the end of the coding region after the 5' untranslated region. This is backwards and cannot be used.

To isolate the mutated *e(r)* gene comprised of the 5' and 3' untranslated regions and the human coding region, an EcoRI-XhoI double digest was done. These are the two recognition sites flanking either end of the untranslated regions within the pSmart plasmid (Fig. 7). The gene alone is about 3.0 kb. After running the digest on a gel, an extraction and purification of the band at 3.0 kb was done to completely isolate the *e(r)* gene.

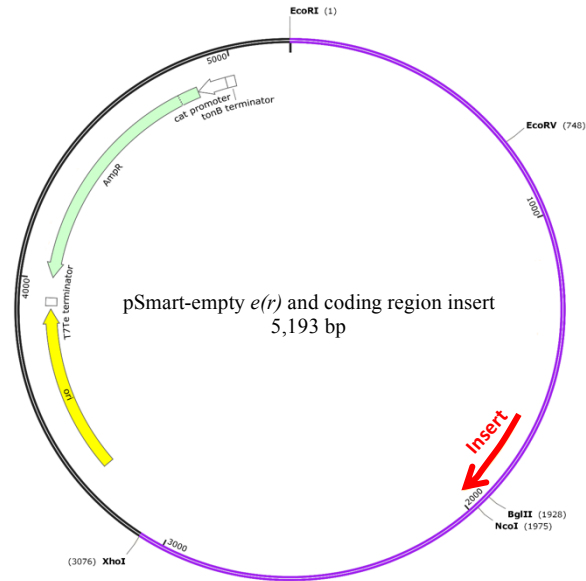


Figure 7. pSmart-empty *e(r)* and Coding Region Insert

The pSmart vector complete with the 5' and 3' untranslated regions and the mutated human coding region in between (purple). The coding region insert is represented by a red arrow. Important restriction enzyme sites are labeled.

3. Insertion of the Mutated Human *e(r)* Gene into the Final Vector, pattB

A relatively new system called ϕ C31 is the chosen method to insert the mutated gene into *Drosophila*. This choice is essential to the phenotypic study and conclusions made following insertion into *Drosophila* embryos because it removes the possibility that the new gene is in a different location of the genome in each fly.

The method that has been used previously for these types of studies utilized P transposable elements which insert randomly into the genome (Rubin and Spradling 1982). This has caused ambiguity in the results because it is hard to verify whether observations are due to the different chromosome locations or the actual mutation that is being studied. In contrast, the ϕ C31 system uses ϕ C31 integrase, isolated from *Streptomyces* temperate bacteriophages, in order to perform site-specific insertion (Thorpe and Smith 1998). ϕ C31 Integrase catalyzes a reaction that causes two attachment sites to recombine.

In order to use this system in *Drosophila* genetic studies, one attachment site (attP) was put into the genome using transposons. A stock of flies is then purified so that every single one has the attachment site in the same location. Then, when a plasmid is introduced that also

contains an attachment site (attB) along with ϕ C31 integrase, they can integrate, ensuring that the new gene is in the same place in the genome of each resulting embryo (Groth et al. 2004).

In this specific study, the isolated 3.0 kb mutated human *e(r)* gene was inserted into a plasmid referred to as pattB that was cut with EcoRI and XhoI (Fig. 8). After verifying that this insertion was successful, the purified pattB vector with the 3.0 kb insert was sent to BestGene Inc (<https://www.thebestgene.com>) where they perform the ϕ C31 system. They inject ϕ C31 integrase and the pattB vector sent to them into embryos. Next, they select for the transformants (flies that contain the integrated transgene). The transformed stocks are then sent back to Dr. Tsubota.

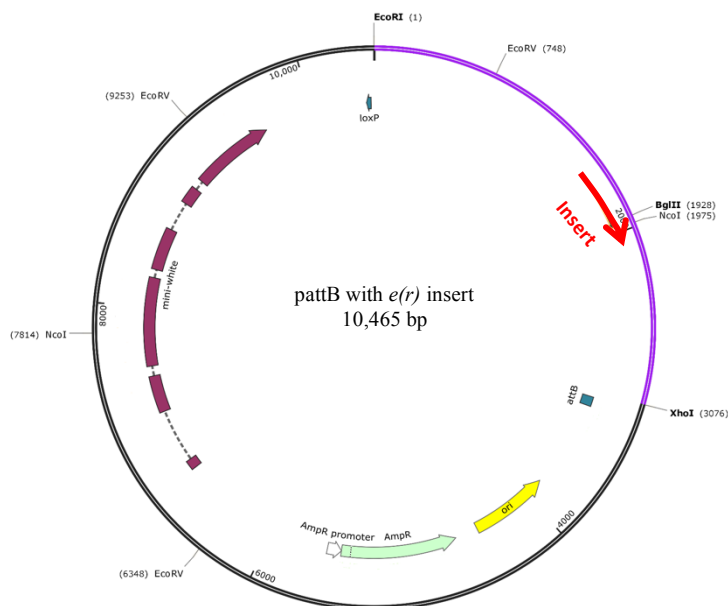


Figure 8. pattB Vector with *e(r)* Gene Insert

The pattB vector with the inserted 3.0 kb *e(r)* gene (purple) with a mutation in the coding region (red arrow). The attachment site (attB) is labeled as well as the mini-white gene that will help in identifying the organisms that contain the plasmid. Important restriction enzyme sites are labeled as well.

4. Phenotypic Analysis of Mutant *Drosophila Melanogaster* Using Genetic Manipulation

An important aspect of the pattB vector is the mini-white gene that will cause the *Drosophila* containing the vector to have red eyes. Once the flies from BestGene Inc. with the vector in their gametes reproduce, those that contain the plasmid in their chromosomes will have red eyes, making them identifiable from those that do not (which have white eyes).

To set up crosses that will be used to analyze the effect of the various mutations in the *e(r)* gene, one male with red eyes, thereby containing the transgene, is put with females from a heterozygous stock of one of the tested phenotypic mutations. It is important to maintain a heterozygous stock because if the *Drosophila* have an *e(r)* deletion in both sex chromosomes, then they will not survive. However, if the deletion is in only one, then the wild-type gene is dominant and the organism survives.

A Punnett square of the cross is shown in Fig. 9. The males used in the cross have a normal X and Y chromosome but the *e(r)* transgene on both copies of an autosome. The transgene is capable of rescuing if a male progeny with an *e(r)* deletion in their only X chromosome results in a functional ERH protein, thereby proving that the mutated copy of *e(r)* on an autosome is functional. If the *Drosophila* males are in fact rescued by the transgene and survive, they will have red eyes. The female progeny will always survive since at least one of their X chromosomes will be from the male containing a normal and functional *e(r)* gene.

In normal reproduction, there should theoretically be a 1:1 ratio of males to females. In this cross, all females that inherit the deletion will survive. By counting the number of these females, identified by red eyes, they can be compared to the red-eyed males that are viable. The ratio of red-eyed males to females can reveal whether a mutation rescued an *e(r)* deletion. This is the premise of the analyses to determine the viability of each mutation.

		Heterozygous Female Stock	
		$X^{e(r)-}; +$	$X^{+}; +$
Male with Mutated <i>e(r)</i> gene	$X^{+}; e(r)$ transgene	Female 1: $X^{e(r)-}/X^{+};$ $+/e(r)$ transgene	Female 2: $X^{+}/X^{+};$ $+/e(r)$ transgene
	$Y; e(r)$ transgene	Male 1: $X^{e(r)-}/Y;$ $+/e(r)$ transgene	Male 2: $X^{+}/Y;$ $+/e(r)$ transgene

Figure 9. Cross of Heterozygous Female with *e(r)* Deletion and Male with *e(r)* Transgene

The cross between a heterozygous stock of females with a male that contains a mutation to be tested. The Punnett square shows the gametes that could come together to create different offspring. The contents of the gametes and resulting offspring are shown in the following format: sex chromosome(s); autosome allele(s). Female 1 and Male 1 are the ones of interest when determining whether each mutation can rescue an *e(r)* deletion. They can be identified by their red eyes, caused by the mini-white gene incorporated in the pattB vector, as opposed to the white eyed flies with the wildtype *e(r)* gene on all sex chromosomes. Note that this is a generalized

example that can be applied to each mutant phenotype that was tested ($e(r)^{27-1}$ r^{hd1} double mutant, $e(r)^{37-6}$ N^{nd-p} double mutant, $e(r)^{27-1}$ single mutant)

Results

1. Evidence of the Successful Creation of Recombinant Vectors

When creating the recombinant vectors containing the mutated human coding region within the pSmart vector, there were two different ways that the insert could go into the vector. This was due to the fact that both ends of the coding region matched both sticky ends of the vector. In order to determine whether the desired orientation was achieved, a restriction enzyme digest was done to differentiate between the two. The enzymes NcoI and EcoRV would create the following sized fragments in the two possible cases:

- 1) Incorrect Orientation: 0.9 kb and 4.2 kb
- 2) Desired Orientation: 1.2 kb and 3.9 kb

Following the NcoI-EcoRV digest, gel electrophoresis was used to identify the correctly orientated cultures. Fig. 10 shows the two cases and evidence that the plasmids used to complete the study were in the proper orientation.

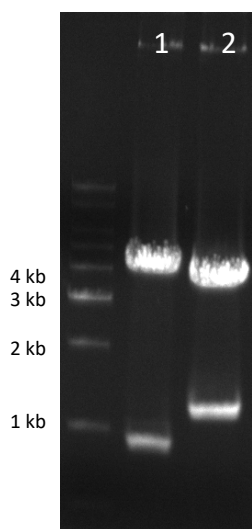


Figure 10. Gel Electrophoresis of pSmart Vector with Coding Region Insert

An example of how the correct orientation of the insert was identified. The plasmid contains the human coding region with the mutation Y92F. Lane 1 is not the desired orientation as the lower band is below 1 kb, presumably at about 0.9 kb. Lane 2 is the correct orientation as the upper band is at approximately 3.9 kb and the lower band is just over 1 kb or approximately 1.2 kb.

After the extraction of the $e(r)$ gene, complete with both untranslated regions and the manipulated human coding region, verification of purity and size was necessary. Without proper

isolation, the rest of the study would be compromised. The expected band sizes were the 3.0 kb fragments of the *e(r)* gene and the remaining 2.1 kb of the vector. Evidence of successful isolation is in Fig. 11.

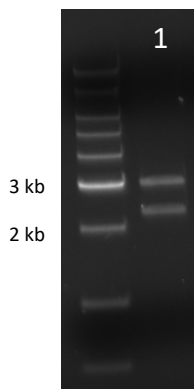


Figure 11. Gel Electrophoresis of Isolated *e(r)* Gene

An example of successful isolation of the *e(r)* gene. At about 3.0 kb is the fragment that was isolated for the eventual insertion into the pattB vector.

As seen in Fig. 8, once the *e(r)* gene is inserted into the vector referred to as pattB, the entire plasmid is approximately 10.4 kb. A digest with the enzymes XhoI and XbaI can be used to verify the total size. The expected sizes of the fragments are 2.0 kb and 8.4 kb. Evidence of successful insertion of the *e(r)* gene is in Fig. 12.

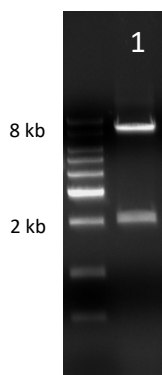


Figure 12. Digest of pattB Plasmid and *e(r)* Gene Insert for Verification

An example of a digested pattB plasmid containing the insert. The gel is evidence that the vector is of correct size and can be inserted into the *Drosophila* genome.

2. Results of the Crosses of Mutant Females to Males Containing the Designed Transgenes

The cross described in Fig. 9 was done with each *e(r)* mutant and all three of the *e(r)* mutant test stocks. The results of the *e(r)*²⁷⁻¹ single mutant cross with each tested change is shown in Table 1. The subsequent crosses of both the *e(r)*²⁷⁻¹ *r^{hd1}* double mutant and *e(r)*³⁷⁻⁶ *N^{md-p}*

double mutant are shown in Tables 2 and 3. For all three tests, a control was done in order to find the ratio of males to females that should be expected. This was done with the same cross but the transgene did not have any mutations and was the wild-type human $e(r)$ gene.

Table 1. The $e(r)^{27-1}$ Mutation Results

Transgene	Number of Males ($X^{e(r) 27-1} / Y$; +/e(r) transgene)	Number of Females ($X^{e(r) 27-1} / X^{+}$; +/e(r) transgene)	Ratio of Males:Females
Control	245	278	0.881
T18E S24A	461	555	0.831
T18A S24A	184	209	0.880
Y19F	204	229	0.891
Y22F	292	278	1.05
S47A	371	388	0.956
S47E	304	302	1.01
Y92F	212	220	0.964
K41R	294	274	1.07
K90R	273	240	1.14

Table 1. The cross of each transgene with the $e(r)^{27-1}$ deletion was done and the number of red-eyed males and females was counted. A ratio of males to females was calculated for each. The control test is included.

Table 2. The $e(r)^{27-1} r^{hd1}$ Double Mutant Results

Transgene	Number of Males ($X^{e(r) 27-1 rhd1} / Y$; +/$e(r)$ transgene)	Number of Females ($X^{e(r) 27-1 rhd1} / X^{+}$; +/$e(r)$ transgene)	Ratio of Males:Females
Control	110	297	0.37
T18E S24A	75	434	0.173
T18A S24A	91	250	0.364
Y19F	181	272	0.665
Y22F	98	308	0.318
S47A	129	332	0.389
S47E	90	278	0.324
Y92F	120	278	0.432
K41R	150	268	0.560
K90R	170	201	0.846

Table 2. A cross of each transgene with the $e(r)^{27-1} r^{hd1}$ double mutant was done with the above results. A ratio of males to females was calculated for each and the control test is included for comparison.

Table 3. The $e(r)^{37-6}$ N^{nd-p} Double Mutant Results

Transgene	Number of Males ($X^{e(r) 37-6 N^{nd-p}}/Y$; +/e(r) transgene)	Number of Females ($X^{e(r) 37-6 N^{nd-p}}/X^{+}$; +/e(r) transgene)	Ratio of Males:Females
Control	99	391	0.253
T18E S24A	4	355	0.011
T18A S24A	14	339	0.041
Y19F	112	318	0.352
Y22F	33	94	0.351
S47A	208	286	0.727
S47E	38	71	0.535
Y92F	94	205	0.459
K41R	183	320	0.572
K90R	228	347	0.657

Table 3. The cross of each transgene with the $e(r)^{37-6}$ N^{nd-p} double mutant was done with the above results. The ratio of males to females was calculated and the control test is included.

Discussion

The following mutations appear to exhibit wild-type results across all three tests: Y19F, Y22F, S47A, S47E, and Y92F. Relative to the ratio of males to females in the controls, they have resulted in rescue. For the changes designed to prevent phosphorylation (Y19F, Y22F, S47A, Y92F), this indicates that phosphorylation at these amino acids is not needed in order for the protein to work properly. Since the mutation designed to mimic phosphorylation (S47E) was also able to rescue, the protein is not inactivated by phosphorylation at serine number 47. Overall, it can be said for position 47 that it has no role in protein activation and function when it comes to phosphorylation since both the prevention and mimic mutations still resulted in wild-type data.

The two mutants designed to test the effects of preventing ubiquitination at sites 41 and 90 resulted in wild-type male to female fly ratios. This means that ubiquitination of either lysine is not required for the activity of ERH. Often times, ubiquitination is a means of marking a protein for degradation. It could be argued that ERH, with one of these mutations, is more stable and is able to remain active longer due to the lack of ubiquitination that may normally occur.

However, at this point the affect on the stability of the protein has not been analyzed. Another possibility is that ubiquitination is deactivating which needs to be studied further.

There are a few possibilities for the results of mutations T18ES24A and T18AS24A. They both had significantly lower viability in the $e(r)^{37-6} N^{nd-p}$ double mutant cross. This could mean that the *Notch* Signaling pathway in particular relies on activity at one of these two sites. In both of the mutations, phosphorylation was prevented at amino acid 24 which suggests that it is most likely the amino acid that needs to be phosphorylated for proper function of the *Notch* pathway. Single mutants need to be synthesized for these two sites in order to verify which is involved in the *Notch* pathway. Another possibility was that mimicked phosphorylation at amino acid 18 was inactivating the *Notch* pathway. However, the T18AS24A results prove that this is not the case as low viability occurred without a mimic in place of threonine 18. There is yet another alternative explanation for the low viability in the $e(r)^{37-6} N^{nd-p}$ double mutant cross. Perhaps the expression of ERH has been decreased too much by altering the amino acids at these positions which is resulting in the low viability that is observed.

When it comes to the test crosses of the T18 and S24 alterations with the $e(r)^{27-1} r^{hdl}$ double mutant, it is hard to compare the results. This is because the double mutant, which was thought to be completely lethal, resulted in the presence of males with white eyes during a later round of crosses. This involved the T18AS24A mutant cross. The presence of males surviving the “lethal” double mutation adds an extra internal control to those crosses that the others do not have. Though the T18ES24A cross with $e(r)^{27-1} r^{hdl}$ has almost half the ratio of T18AS24A, they cannot be compared properly due to the differences in control.

Conclusion

The mutations that have been crossed with all three mutant phenotypes have proven to have no effect on the activity of ERH with the exception of T18ES24A and T18AS24A. It has been concluded that the serine at position 47 does not have any role in protein activation as the prevention or mimic of phosphorylation has shown to produce wild-type results. The tyrosines at positions 19, 22 and 92 have also rescued the mutant phenotypes when prevented from phosphorylation. However, another study should be done to test the possibility that phosphorylation can inactivate ERH. Both of the double amino acid changes at sites 18 and 24 have shown decreased viability from the wild-type in the $e(r)^{37-6} N^{nd-p}$ cross. Reasons for this observation may include an interference with the *Notch* signaling pathway or an overall decrease

in protein activity. Although some of these potential phosphorylation sites have proven to have no role in protein activation, there may be other reasons for their conservation across species that could be further explored in subsequent studies. One potential possibility that can be explored in the future is that the amino acids T18, Y19, Y22, and S24 have a redundant function. Since they are so close together, it is possible that the phosphorylation of any one of them is enough for protein activation. A quadruple mutant with T18A, Y19F, Y22F, and S24A could reveal whether this is true in another trial.

Appendices

- I. Human coding region sequences sent to Integrated DNA Technologies, Inc. to be synthesized. The coding portion of the DNA sequences starts at the double-underlined ATG and ends at the stop codon TAA which is before the double-underlined portion at the end. The underlined sequence at the beginning notes the PciI site and the underlined sequence at the end notes the NcoI site. Within each sequence is an underlined codon change for that particular mutant. The corresponding polypeptide sequence has the amino acid change(s) underlined as well.

Phosphorylation Sites

Drosophilized Human *e(r)* **S47A**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCACCTACGCGGACTAC
GAGAGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGAAGCGCATGAATCCAAACGCCCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG
CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGAAGATCTATGTGCTGCTCCGTGCGCCAGGCA
CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT
MSHTILLVQPTKRPEGRTYADYESVNECMEGVCKMYEEHLKRMNPNAPSITYDISQLFDFIDDLADLSCL
VYRADTQTYQPYNKDWIKEKIYVLLRRQAQAGK

Drosophilized Human *e(r)* **S47E**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCACCTACGCGGACTAC
GAGAGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGAAGCGCATGAATCCAAACGAGCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG
CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGAAGATCTATGTGCTGCTCCGTGCGCCAGGCA
CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT
MSHTILLVQPTKRPEGRTYADYESVNECMEGVCKMYEEHLKRMNPNEPSITYDISQLFDFIDDLADLSCL
VYRADTQTYQPYNKDWIKEKIYVLLRRQAQAGK

Drosophilized Human *e(r)* **Y92F**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCACCTACGCGGACTAC
GAGAGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGAAGCGCATGAATCCAAACAGCCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG
CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGAAGATCTTTCGTGCTGCTCCGTGCGCCAGGCA
CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT

MSHTILLVQPTKRPEGRITYADYESVNECMGVCKMYEEHLKRMNPNSPSITYDISQLFDFIDDLADLSCL
VYRADTQTYQPYNKDWIKEKIFVLLRRQAQQAGK

Drosophilized Human e(r) **Y22F**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCACCTACGCGGACTTTC
GAGAGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGAAGCGCATGAATCCAAACAGCCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG
CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGAAGATCTATGTGCTGCTCCGTCGCCAGGCA
CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT

MSHTILLVQPTKRPEGRITYADYESVNECMGVCKMYEEHLKRMNPNSPSITYDISQLFDFIDDLADLSCL
VYRADTQTYQPYNKDWIKEKIYVLLRRQAQQAGK

Drosophilized Human e(r) **Y19F**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCACCCTTCGCGGACTAC
GAGAGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGAAGCGCATGAATCCAAACAGCCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG
CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGAAGATCTATGTGCTGCTCCGTCGCCAGGCA
CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT

MSHTILLVQPTKRPEGRITYADYESVNECMGVCKMYEEHLKRMNPNSPSITYDISQLFDFIDDLADLSCL
VYRADTQTYQPYNKDWIKEKIYVLLRRQAQQAGK

Drosophilized Human e(r) **T18A S24A**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCCTTACGCGGACTAC
GAGCGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGAAGCGCATGAATCCAAACAGCCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG
CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGAAGATCTATGTGCTGCTCCGTCGCCAGGCA
CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT

MSHTILLVQPTKRPEGRITYADYESVNECMGVCKMYEEHLKRMNPNSPSITYDISQLFDFIDDLADLSCL
VYRADTQTYQPYNKDWIKEKIYVLLRRQAQQAGK

Drosophilized Human e(r) **T18E S24A**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCCTTACGCGGACTAC
GAGCGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGAAGCGCATGAATCCAAACAGCCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG
CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGAAGATCTATGTGCTGCTCCGTCGCCAGGCA
CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT

MSHTILLVQPTKRPEGRITYADYESVNECMGVCKMYEEHLKRMNPNSPSITYDISQLFDFIDDLADLSCL
VYRADTQTYQPYNKDWIKEKIYVLLRRQAQQAGK

Ubiquitination Sites

Drosophilized Human e(r) **K90R**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCACCTACGCGGACTAC
GAGAGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGAAGCGCATGAATCCAAACAGCCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG
CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGCGCATCTATGTGCTGCTCCGTCGCCAGGCA
CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT

MSHTILLVQPTKRPEGRITYADYESVNECMGVCKMYEEHLKRMNPNSPSITYDISQLFDFIDDLADLSCL
VYRADTQTYQPYNKDWIKEKIYVLLRRQAQQAGK

Drosophilized Human e(r) **K41R**

ACGGAATTCACATGTCGCACACCATCCTATTGGTACAGCCGACCAAGCGTCCAGAGGGCCGCACCTACGCGGACTAC
GAGAGCGTCAACGAGTGCATGGAGGGCGTGTGCAAGATGTACGAGGAGCACCTGCGCGCATGAATCCAAACAGCCC
GAGCATTACGTATGACATTAGCCAGCTATTTCGATTTTCATCGACGATCTGGCGGACCTAAGCTGCCTGGTTTACCGCG

CTGACACCCAGACCTATCAGCCCTACAATAAGGATTGGATCAAGGAGAAGATCTATGTGCTGCTCCGTCGCCAGGCA
 CAGCAGGCTGGCAAGTAACCATGGCTCGAGGAT
 MSH TILLVQPTKRPEGR TYADYESVNECM EGVCKMYEEHL ERMNPNSPSITYDISQLFDFIDDLADLSCL
 VYRADTQTYQPYNKDWIKEKIYVLLRRQAQAGK

Bibliography

Arai R, Kukimoto-Niino M, Uda-Tochio H, Morita S, Uchikubo-Kamo T, Akasaka R, Etou Y, Hayashizaki Y, Kigawa T, Terada T, et al. 2005. Crystal structure of an enhancer of rudimentary homolog (ERH) at 2.1 Å resolution. *Protein Sci Publ Protein Soc.* 14:1888–1893. doi:10.1110/ps.051484505.

Gelsthorpe M, Pulumati M, McCallum C, Dang-Vu K, Tsubota SI. 1997. The putative cell cycle gene, enhancer of rudimentary, encodes a highly conserved protein found in plants and animals. *Gene.* 186:189–195. doi:10.1016/S0378-1119(96)00701-9.

Gelsthorpe ME, Tan Z, Phillips A, Eissenberg JC, Miller A, Wallace J, Tsubota SI. 2006. Regulation of the *Drosophila melanogaster* Protein, Enhancer of Rudimentary, by Casein Kinase II. *Genetics.* 174:265–270. doi:10.1534/genetics.106.061465.

Groth AC, Fish M, Nusse R, Calos MP. 2004. Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics.* 166:1775–1782.

Rubin GM, Spradling AC. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science.* 218:348–353.

Thorpe HM, Smith MCM. 1998. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci.* 95:5505–5510.

Tsubota S, Ryan T, Rizzo N, Hing H. 2016. The human and *Drosophila* ERH are functionally equivalent: Evidence from transgenic studies. *Intern Med Rev.* 2. doi:10.18103/imr.v2i8.154.

Tsubota SI, Vogel AC, Phillips AC, Ibach SM, Weber NK, Kostrzebski MA, Spencer SA. 2011. Interactions between enhancer of rudimentary and Notch and deltex reveal a regulatory function of enhancer of rudimentary in the Notch signaling pathway in *Drosophila melanogaster*. *Fly (Austin).* 5:275–284. doi:10.4161/fly.5.4.17807.

Vicario S, Moriyama EN, Powell JR. 2007. Codon usage in twelve species of *Drosophila*. *BMC Evol Biol.* 7:226. doi:10.1186/1471-2148-7-226.

Wojcik E, Murphy AM, Fares H, Dang-Vu K, Tsubota SI. 1994. Enhancer of Rudimentary(p1), E(r)(p1), a Highly Conserved Enhancer of the Rudimentary Gene. *Genetics.* 138:1163–1170.

Xu G, Jaffrey SR. 2013. Proteomic Identification of Protein Ubiquitination Events. *Biotechnol Genet Eng Rev.* 29:73–109. doi:10.1080/02648725.2013.801232.

Zafrakas M, Losen I, Knüchel R, Dahl E. 2008. Enhancer of the rudimentary gene homologue (ERH) expression pattern in sporadic human breast cancer and normal breast tissue., Enhancer of the rudimentary gene homologue (ERH) expression pattern in sporadic human breast cancer and normal breast tissue. BMC Cancer BMC Cancer. 8, 8:145, 145–145. doi:10.1186/1471-2407-8-145, 10.1186/1471-2407-8-145.